

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Shigeo Yoshida et al.

Art Unit : Unknown

Serial No. :

Examiner : Unknown

Filed : October 26, 2001

Title : GENOMIC DNA ANALYSIS PROGRAM

#5
10/26/01
10/03/46
10/26/01

Commissioner for Patents
Washington, D.C. 20231

COPY OF PAPERS
ORIGINALLY FILED

TRANSMITTAL OF PRIORITY DOCUMENT UNDER 35 USC §119

Applicant hereby confirms his claim of priority under 35 USC §119 from the following application(s):

· Japan Application No. 2000-327632 filed October 26, 2000

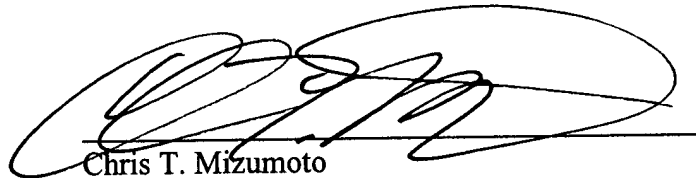
· Japan Application No. 2001-084370 filed March 23, 2001

A certified copy of each application from which priority is claimed is submitted herewith.

Please apply any charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: October 26, 2001



Chris T. Mizumoto
Reg. No. 42,899

Fish & Richardson P.C.
45 Rockefeller Plaza, Suite 2800
New York, New York 10111
Telephone: (212) 765-5070
Facsimile: (212) 258-2291

30070524.doc

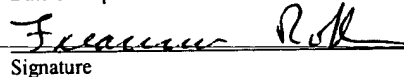
CERTIFICATE OF MAILING BY EXPRESS MAIL

Express Mail Label No. EF045065430US

I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as Express Mail Post Office to Addressee with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, Washington, D.C. 20231.

October 26, 2001

Date of Deposit



Signature

Francisco Robles

Typed or Printed Name of Person Signing Certificate

(Translation)

1c872 U.S. P.
10/003446
10/26/01

PATENT OFFICE
JAPANESE GOVERNMENT

COPY OF PAPERS
ORIGINALLY FILED

This is to certify that the annexed is a true copy of the
following application as filed with this Office.

Date of Application: March 23, 2001
Application Number: Japanese Patent Application
No. 084370/2001
Applicant(s): RIKEN

August 3, 2001

Commissioner,
Patent Office

Kozo Oikawa (seal)

Certificate No. 2001-3069888

日 本 国 特 許 庁
JAPAN PATENT OFFICE

JC872 U.S. PRO
10/003446
10/26/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 3月23日

出 願 番 号

Application Number:

特願2001-084370

出 願 人

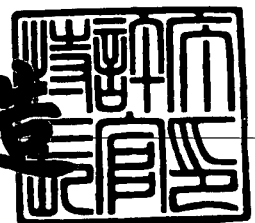
Applicant(s):

理化学研究所

2001年 8月 3日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3069888

【書類名】 特許願

【整理番号】 RJH12-138

【提出日】 平成13年 3月23日

【あて先】 特許庁長官 殿

【国際特許分類】 G01N 33/48

【発明の名称】 ゲノムDNAの解析プログラムを記録した記録媒体

【請求項の数】 17

【発明者】

 【住所又は居所】 埼玉県和光市広沢2番1号 理化学研究所内

 【氏名】 吉田 茂男

【発明者】

 【住所又は居所】 埼玉県和光市広沢2番1号 理化学研究所内

 【氏名】 松山 知樹

【発明者】

 【住所又は居所】 埼玉県和光市広沢2番1号 理化学研究所内

 【氏名】 阿部 知子

【発明者】

 【住所又は居所】 埼玉県和光市広沢2番1号 理化学研究所内

 【氏名】 戎崎 俊一

【特許出願人】

 【識別番号】 000006792

 【氏名又は名称】 理化学研究所

【代理人】

 【識別番号】 100091096

 【弁理士】

 【氏名又は名称】 平木 祐輔

【選任した代理人】

 【識別番号】 100096183

 【弁理士】

【氏名又は名称】 石井 貞次

【先の出願に基づく優先権主張】

【出願番号】 特願2000-327632

【出願日】 平成12年10月26日

【手数料の表示】

【予納台帳番号】 015244

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9503608

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ゲノムDNAの解析プログラムを記録した記録媒体

【特許請求の範囲】

【請求項1】 コンピュータに、

ゲノム塩基配列情報に基づいて、コントロールの二次元電気泳動パターンを作成する手順と、

上記コントロールの二次元電気泳動パターンと、解析対象のゲノムDNAを用いて二次元電気泳動を行って得られた解析対象の二次元電気泳動パターンとを比較する手順とを含み、

上記コントロールの二次元電気泳動パターンと上記解析対象の二次元電気泳動パターンとにおけるスポット位置の差異を検出させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項2】 上記コントロールの二次元電気泳動パターンを作成する手順は、上記ゲノム塩基配列情報における第一の制限酵素の認識配列及び第二の制限酵素の認識配列を検出し、当該第一の制限酵素の認識配列及び当該第二の制限酵素の認識配列により挟まれた塩基配列と、当該第一の制限酵素の認識配列及び当該第一の制限酵素の認識配列より挟まれた塩基配列とに基づいて行うことを特徴とする請求項1記載の記録媒体。

【請求項3】 上記第一の制限酵素及び第二の制限酵素は、メチル化非感受性制限酵素から選ばれることを特徴とする請求項2記載の記録媒体。

【請求項4】 上記ゲノム塩基配列情報を、通信回線網を介して取得する手順を含むことを特徴とする請求項1記載の記録媒体。

【請求項5】 上記コントロールの二次元電気泳動パターンを作成する手順では、ゲノム塩基配列情報に基づいて複数のスポットを作成するとともにこれら複数のスポットをゲノム上での遺伝子座情報とリンクさせることを特徴とする請求項1記載の記録媒体。

【請求項6】 解析対象のゲノムDNAを用いて二次元電気泳動を行って二次元電気泳動パターンを作成する工程と、

上記二次元電気泳動パターンと、ゲノム塩基配列情報に基づいて作成したコン

トロールの二次元電気泳動パターンとを比較する工程とを含み、

上記コントロールの二次元電気泳動パターンと上記解析対象の二次元電気泳動パターンとにおけるスポット位置の差異を検出することを特徴とするゲノムDNA解析方法。

【請求項7】 上記解析対象のゲノムDNAを植物細胞から抽出する工程を含むことを特徴とする請求項6記載のゲノムDNA解析方法。

【請求項8】 以下の工程、

- (a) ゲノムDNAを第一の制限酵素で処理する工程、
 - (b) 該制限酵素切断部位に標識を付加する工程、
 - (c) 得られたDNA断片を電気泳動により一次元分画を行う工程、
 - (d) 工程(c)により分画されたDNA断片を第二の制限酵素で処理した後、二次元分画を行う工程、
 - (e) 工程(d)により分画された標識DNA断片のスポットを検出する工程、
- によって、上記二次元電気泳動パターンを作成する工程を行うことを特徴とする請求項6記載のゲノムDNA解析方法。

【請求項9】 上記工程(b)の後、第一の制限酵素及び第二の制限酵素と異なる制限酵素により処理した後に上記工程(c)を行うことを特徴とする請求項8記載のゲノムDNA解析方法。

【請求項10】 上記第一の制限酵素の切断部位及び上記第二の制限酵素の切断部位を検出し、これら切断部位に基づいて上記コントロールの二次元電気泳動パターンを作成することを特徴とする請求項8記載のゲノムDNA解析方法。

【請求項11】 上記工程(b)は、上記制限酵素切断部にアダプターの一方を連結するとともに、当該アダプターの他方に上記標識を付加することにより行われることを特徴とする請求項8記載のゲノムDNA解析方法。

【請求項12】 上記コントロールの二次元電気泳動パターンは、ゲノム塩基配列情報に基づいて作成された複数のスポットを有するとともにこれら複数のスポットをゲノム上での遺伝子座情報とリンクさせてなることを特徴とする請求項6記載のゲノムDNA解析方法。

【請求項13】 請求項6～12いずれか一項記載のゲノムDNA解析方法に

より解析対象のゲノムDNAにおける変位箇所を検出することを特徴とする変異遺伝子及び/又は変異形質に関与する遺伝子の同定方法。

【請求項 1 4】 請求項 1 3 に記載の同定方法で検出した変位箇所を含む DNA 断片を、上記二次元電気泳動パターンから単離することを特徴とする変異遺伝子及び/又は変異形質に関与する遺伝子の単離方法。

【請求項 1 5】 ゲノム塩基配列情報に基づいて作成された複数のスポットを有することを特徴とする二次元電気泳動パターン。

【請求項 1 6】 上記複数のスポットは、ゲノム上での遺伝子座情報とリンクされていることを特徴とする請求項 1 5 記載の二次元電気泳動パターン。

【請求項 1 7】 上記ゲノム塩基配列情報は、植物の核ゲノムから得られた塩基配列情報であることを特徴とする請求項 1 5 記載の二次元電気泳動パターン。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、ゲノムDNAを用いた二次元電気泳動パターンを解析するプログラムを記録したコンピュータ読み取り可能な記録媒体に関し、また、二次元電気泳動パターンを用いたゲノムDNA解析方法に関する。

【 0 0 0 2 】

【従来の技術】

近年、高等生物のゲノム解析をより速く行うため、ゲノムスキニング技術の開発が急速に進んでいる。ゲノムスキニングとは、座位（遺伝子座）又はそのコピー数を一度に高速解析し、ゲノムDNAの物理的状态をゲノム全体にわたって検索するものである。ゲノムスキニングによれば、ゲノムの遺伝子地図上における遺伝子の位置、又は1つの遺伝子産物をコードするDNA領域が占める染色体上の区画について、そのシグナルを高速に検出し有無を判定することができる。このゲノムスキニングにおいては、ゲノム上にどのような目印を立てればよいか、ゲノム上のいかなる目印をどのようなシグナルとして検出すればよいか、さらに、より多くの目印をどのように検出し解析すればよいかということが極めて重

要である。これらの目印はランドマークと呼ばれており、ランドマークの染色体上の位置情報を解析することは、ゲノムの連鎖地図又は物理的地図を作製するための大きな基礎的技術となる。従って、ゲノムスキニングを2種類以上のゲノムDNAに適用し比較すると、欠失、増幅、転座などのゲノムDNAの変化を検出することができるようになる。

【0003】

ところで、RLGS (Restriction Landmark Genomic Scanning) 法は、制限酵素で切断したゲノムDNAの単部をランドマークとして標識し、標識されたゲノムDNA断片を二次元電気泳動を用いて展開し、得られるパターンを用いて検出する方法である。このRLGS法は、一度に多量のランドマークを検出することができるため、ゲノムDNAの変化を高速に検出できる技術であるといえる。

【0004】

このRLGS法には、解析対象のゲノムDNAから得られた二次元電気泳動パターンと比較するため、コントロールとなる二次元電気泳動パターンを必要とする。コントロールとなる二次元電気泳動パターンは、野生型の細胞から抽出したゲノムDNAを用いて同様に二次元電気泳動を行うことによって得ることができる。RLGS法では、コントロールの二次元電気泳動パターンを得るための工程が煩わしく、また、得られたコントロールの二次元電気泳動パターンを用いた比較も煩わしい。さらに、スポットのクローニングには、著しい労力と煩雑な技術を必要とするといった問題点があった。

【0005】

【発明が解決しようとする課題】

そこで、本発明は、比較の対象となるコントロールの二次元電気泳動パターンを容易に、且つ、様々な条件で作成することができるゲノムDNAを用いた二次元電気泳動パターンを解析するプログラムを記録したコンピュータ読み取り可能な記録媒体を提供することを目的とする。また、本発明は、コントロールの二次元電気泳動パターンを容易に、且つ、様々な条件で作成することによって、迅速且つ確実な解析を行うことができるゲノムDNA解析方法を提供することを目的とする。

【0006】

【課題を解決するための手段】

(1) 上述した目的を達成した本発明は、コンピュータに、ゲノム塩基配列情報に基づいて、コントロールの二次元電気泳動パターンを作成する手順と、上記コントロールの二次元電気泳動パターンと、解析対象のゲノムDNAを用いて二次元電気泳動を行って得られた解析対象の二次元電気泳動パターンとを比較する手順とを含み、上記コントロールの二次元電気泳動パターンと上記解析対象の二次元電気泳動パターンとにおけるスポット位置の差異を検出させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体である。

【0007】

(2) また、本発明は、上記コントロールの二次元電気泳動パターンを作成する手順は、上記ゲノム塩基配列情報における第一の制限酵素の認識配列及び第二の制限酵素の認識配列を検出し、当該第一の制限酵素の認識配列及び当該第二の制限酵素の認識配列により挟まれた塩基配列と、当該第一の制限酵素の認識配列及び当該第一の制限酵素の認識配列より挟まれた塩基配列とに基づいて行うことを特徴とする(1)記載の記録媒体である。

【0008】

(3) さらに、本発明は、上記第一の制限酵素及び第二の制限酵素は、メチル化非感受性制限酵素から選ばれることを特徴とする(2)記載の記録媒体である。

(4) さらにまた、本発明は、上記ゲノム塩基配列情報を、通信回線網を介して取得する手順を含むことを特徴とする(1)記載の記録媒体である。

【0009】

(5) さらにまた、本発明は、上記コントロールの二次元電気泳動パターンを作成する手順では、ゲノム塩基配列情報に基づいて複数のスポットを作成するとともにこれら複数のスポットをゲノム上での遺伝子座情報とリンクさせることを特徴とする(1)記載の記録媒体である。

【0010】

(6) さらにまた、本発明は、解析対象のゲノムDNAを用いて二次元電気泳動

を行って二次元電気泳動パターンを作成する工程と、上記二次元電気泳動パターンとゲノム塩基配列情報に基づいて作成したコントロールの二次元電気泳動パターンとを比較する工程とを含み、上記コントロールの二次元電気泳動パターンと上記解析対象の二次元電気泳動パターンとにおけるスポット位置の差異を検出することを特徴とするゲノムDNA解析方法である。

【 0 0 1 1 】

(7) さらにまた、本発明は、上記解析対象のゲノムDNAを植物細胞から抽出する工程を含むことを特徴とする(6)記載のゲノムDNA解析方法である。

(8) さらにまた、本発明は、以下の工程、

- (a) ゲノムDNAを第一の制限酵素で処理する工程、
- (b) 該制限酵素切断部位に標識を付加する工程、
- (c) 得られたDNA断片を電気泳動により一次元分画を行う工程、
- (d) 工程(c)により分画されたDNA断片を第二の制限酵素で処理した後、二次元分画を行う工程、
- (e) 工程(d)により分画された標識DNA断片のスポットを検出する工程、

によって、上記二次元電気泳動パターンを作成する工程を行うことを特徴とする(6)記載のゲノムDNA解析方法である。

【 0 0 1 2 】

(9) さらにまた、本発明は、上記工程(b)の後、第一の制限酵素及び第二の制限酵素と異なる制限酵素により処理した後に上記工程(c)を行うことを特徴とする(8)記載のゲノムDNA解析方法である。

(10) さらにまた、本発明は、上記第一の制限酵素の切断部位及び上記第二の制限酵素の切断部位を検出し、これら切断部位に基づいて上記コントロールの二次元電気泳動パターンを作成することを特徴とする(8)記載のゲノムDNA解析方法である。

【 0 0 1 3 】

(11) さらにまた、本発明は、上記工程(b)は、上記制限酵素切断部にアダプター的一方を連結するとともに、当該アダプターの他方に上記標識を付加することにより行われることを特徴とする(8)記載のゲノムDNA解析方法である

(12) さらにまた、本発明は、上記コントロールの二次元電気泳動パターンは、ゲノム塩基配列情報に基づいて作成された複数のスポットを有するとともにこれら複数のスポットをゲノム上での遺伝子座情報とリンクさせてなることを特徴とする(6)記載のゲノムDNA解析方法である。

【0014】

(13) さらにまた、本発明は、(6)～(12)いずれか一項記載のゲノムDNA解析方法により解析対象のゲノムDNAにおける変位箇所を検出することを特徴とする変異遺伝子及び/又は変異形質に関与する遺伝子の同定方法である。

(14) さらにまた、本発明は、請求項13に記載の同定方法で検出した変位箇所を含むDNA断片を、上記二次元電気泳動パターンから単離することを特徴とする変異遺伝子及び/又は変異形質に関与する遺伝子の単離方法である。

【0015】

(15) ゲノム塩基配列情報に基づいて作成された複数のスポットを有することを特徴とする2次元電気泳動パターン。

(16) 上記複数のスポットは、ゲノム上での遺伝子座情報とリンクされていることを特徴とする(15)記載の2次元電気泳動パターン。

(17) 上記ゲノム塩基配列情報は、植物の核ゲノムから得られた塩基配列情報であることを特徴とする(15)記載の2次元電気泳動パターン。

【0016】

【発明の実施の形態】

以下、本発明を詳細に説明する。

本発明に係る記録媒体は、以下のプログラムを記録し、当該プログラムをコンピュータに実行させることができる。このプログラムは、以下の手順を含む。

【0017】

ゲノム塩基配列情報に基づいて、コントロールの二次元電気泳動パターンを作成する手順(以下、手順1)。

上記コントロールの二次元電気泳動パターンと、解析対象のゲノムDNAを用いて二次元電気泳動を行って得られた二次元電気泳動パターンとを比較する手順(

以下、手順2)。

【0018】

すなわち、このプログラムは、解析対象から抽出したゲノムDNAから得られる二次元電気泳動パターンと、コントロールとなる二次元電気泳動パターンとを、コンピュータを用いて比較するものである。このプログラムにおいて、手順1では、先ず、コントロールとなる二次元電気泳動パターンを得る。手順2では、別途、得られた解析対象の二次元電気泳動パターンを用いる。

【0019】

先ず、解析対象となるゲノムDNAの二次元電気泳動パターンを得る方法について説明する。二次元電気泳動パターンは、制限酵素の認識部位を目印（ランドマーク）として用い、このランドマークをシグナルとして検出する方法（Restriction Landmark Genomic Scanning法（RLGS法）；Hatada, I. et al., Proc. Natl. Acad. Sci., USA, 88, 9523-9527, 1991）を基本概念として適用する。植物のゲノムDNAの解析に関しては、Matsuyama et al. Plant. Mol. Biol. Rep. 18 : 331-338, 2000を基本理念として適用する。解析対象となるゲノムDNAの二次元電気泳動パターンは、以下の方法に準じて得ることができる。

【0020】

(1) ゲノムDNAの第一の制限酵素による処理

本発明の解析方法の対象となるゲノムDNAは特に限定されるものではなく、植物由来のものであると他の生物種（動物、細菌、酵母等）のものであるとを問わない。植物由来のゲノムDNAとしては、例えばイネ、タバコ、アラビドプシス、コムギ、トマト等由来のものが挙げられるが、イネ又はアラビドプシス由来のものが好ましい。

【0021】

動物由来のゲノムDNAとしては、例えばヒト、マウス、センチュウ等由来のものが挙げられる。細菌由来のゲノムDNAとしては、例えば枯草菌、大腸菌等由来のものが挙げられる。

本発明に係るプログラム及びDNA解析方法においては、特に、上述した植物由来のゲノムDNAを解析対象とすることが好ましい。植物由来のゲノムDNAは、従来

より公知の方法 (Dilaport et al. Plant Mol. Biol. Rep. 1: 19-21, 1983) に準じて、植物に対して除タンパク質処理、除多糖質処理等を施すことによって得ることができる。

【0022】

この「ゲノムDNAの第一の制限酵素による処理」では、先ず、抽出したゲノムDNAを第一の制限酵素 (図1(1)において「A」の位置を切断する酵素であり、制限酵素Aとする) で切断する。制限酵素Aとしては、切断したときに生じる平均断片長が100kbを超える程度のもの、すなわち平均100kbを超える間隔でしか存在しない制限酵素部位を認識する制限酵素であって6~8塩基認識のいわゆるレアカッター制限酵素が挙げられる。

【0023】

上記制限酵素Aは、切断したときに生じる制限酵素部位の5'側が突出した付着末端となるように切断するもの (5'突出型という)、及び3'側が突出した付着末端となるように切断するもの (3'突出型という) のいずれを用いることもできる。さらに、突出部の長さは1塩基以上のものが挙げられ、十分なシグナル強度を得るためには、突出部の長さが2塩基以上となるように切断する酵素が好ましい。このような制限酵素Aとしては、5'突出型については例えばNotI、BssHII、AccIIIが挙げられ、3'突出型については例えばBstXI、BglI、MwoIが挙げられる。

【0024】

ここで、特に解析対象を植物細胞由来のゲノムDNAとする場合、第一の制限酵素としては、メチル化非感受性の制限酵素を使用することが好ましい。メチル化非感受性の制限酵素を使用することによって、植物細胞由来のゲノムDNAのようにメチル化されたシトシンを有するゲノムDNAであっても確実に切断することができる。メチル化非感受性の制限酵素としては、例えば、AccIII及びPacI等を挙げることができる。

【0025】

(2) 制限酵素切断部位への標識

次に、上記制限酵素Aの切断部位に対して標識を付加する。標識を付加する際には、シークエナーゼ (T7DNAポリメラーゼ) により、予め標識された塩基 (標

識用配列)を制限酵素Aの切断部位にフィルイン(fill in; 埋め込むこと)することにより行う。したがって、この場合、シークエナーゼの反応対象となるために、上記制限酵素Aとしては、切断部位の5'側が突出した付着末端となるように切断するもの(5'突出型という)を使用することが好ましい。

【0026】

標識物質としては、 $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ 、 $[\alpha\text{-}^{32}\text{P}]\text{dGTP}$ 等の放射性同位体、テトラメチル-ローダミン-6-d UTP、フルオレセイン-12-d UTP等の蛍光色素等が挙げられ、任意に選択することができる。

また、標識を付加する際には、制限酵素Aの切断部位にアダプターを付加し、当該アダプターに標識を付加してもよい。アダプターは、その一方の側を上記制限酵素Aの切断部位に連結することができる配列(連結配列という)として設計し、他方の側はアダプターを標識するための配列(標識用配列という)として設計する。アダプターの標識用配列を5'側が突出した付着末端とすることによって、シークエナーゼによる反応対象とすることができ、標識用配列に対して標識を付加することができる。

【0027】

このようにアダプターを介して標識を付加する場合、制限酵素Aとしては、切断部位の3'側が突出した付着末端となるように切断するもの、すなわち、3'突出型を使用することが好ましい。この場合、制限酵素Aによる切断部位には、シークエナーゼ(T7DNAポリメラーゼ)によるフィルインを防止することができる。言い換えると、アダプターの連結配列は、5'側が突出した付着末端とすることが好ましい。

【0028】

ここで、連結配列及び標識用配列は一本鎖で構成される。また、連結配列と標識用配列との間は5~45塩基からなる二本鎖で構成され、クローニングのしやすさを考慮すると、20~35塩基からなるものが好ましい。連結配列は、上記制限酵素Aの認識配列に対応して設計することができる。また、標識用配列及び二本鎖領域の配列は任意に設計することができる。

【0029】

標識用配列の長さは任意に設定することができるので、標識用配列の長さが長いほど検出感度を高めることができる。但し、標識用配列が長すぎるとDNAの二次構造（例えばステムループ構造又はヘアピン構造）が形成されることがある。そこで、これらの二次構造を形成しないよう、標識用配列は2～10塩基であることが好ましい。

【0030】

上述したように、制限酵素Aの切断部位に標識を直接付加する場合には、タバコ、アラビドプシス等に対するスポット検出能が低いため、全てのゲノムDNAに適用することが困難である。そこで、アダプターを介して標識を付加することにより、全てのゲノムDNAに適用することができ、検出能を改善することができる。

【0031】

また、アダプターを使用する場合、制限酵素Aとしては、認識配列中にN(NはA、G、C又はTを表す)を有するBstXIを用いても良い。BstXIのように3'突出型であれば、後述する標識の際にシークエナーゼ（T7DNAポリメラーゼ）の反応対象とならず、標的DNA断片以外の断片を標識することから防ぐことができる。なお、このような3'突出型で2塩基以上の突出部を有する酵素はBstXI、BglI及びMwoIのほかにも多数存在し（例えばAlwNI、BanII、BsiEI、BsiHKAI、BslI、BsmI、BspI286I、BsrDI、DraIII、DrdI、PflMI、SfiI、TspRI、BpmI、BseRI、BsgI、Eco57I）、本発明の方法ではこれらの酵素のいずれも適用することができる。

【0032】

例えば、制限酵素AとしてBstXIを使用すると、BstXIは以下の配列を認識し、3'側が4塩基突出するように（下線部）5'末端から第8番目のNと第9番目のNとの間（*で示す位置）を切断する（図2(1)）。

センス鎖： 5'-CCANNNNN * NTGG-3'（配列番号1）

アンチセンス鎖： 3'-GGTN * NNNNNACC-5'

【0033】

NはA、G、C及びTのいずれでもよいので、アダプターの突出部の4つの配列は全ての組み合わせ（ $4^4=256$ 通り）を考慮して設計される。この場合、ゲノムDNAの

種類や解析の目的（具体的にはアラビドプシス突然変異体における変異領域を検出すること等）に応じて突出部の4塩基を適当な数の組み合わせとなるように規定することができる。例えば、図2(2)にはアダプターの連結配列として「GGGC」が示されており、この場合は、アダプターはBstXIによる切断断片のうち「CCCG」配列を突出配列とする1種類の断片（すなわち突出部が「CCCG」のみの配列を有する断片）と連結することとなる（図2(3)）。同様に、アダプターの連結配列を「GGGS」（SはG又はCを表す）とした場合は、アダプターはBstXIによる切断断片のうち「CCCG」又は「CCCC」配列を突出配列とする2種類の断片と連結することとなる。従って、BstXIを制限酵素Aとして使用すると、アダプターの連結配列の選択の仕方により1～256通りのランドマーク部位を任意に選択することができる。

なお、アダプターの調製は、通常の化学合成装置（例えばPE Applied Biosystems社のDNA/RNAシンセサイザmodel 394）により得ることができる。そして、アダプターをDNAリガーゼで処理し、ゲノムDNAに連結する。

【0034】

(3) 第二の制限酵素処理

制限酵素Aで切断された断片が100kb以上の長さを有している場合には、そのまま電気泳動等による分画を行うことはできない。そこで、制限酵素Aによる切断断片をさらに短い断片にするため、第二の制限酵素処理を行う（図1(4)）。第二の制限酵素は、切断したときに生じる平均断片長が数～数十kbもの、すなわち平均数～数十kbの間隔で存在する制限酵素部位を認識する制限酵素であって6塩基認識ものを用いることができる（図1においてBの位置を切断する酵素であり、制限酵素Bとする。）。制限酵素Bとしては、例えばEcoRV、DraI等が挙げられる。

なお、制限酵素Aで切断されてなる断片が平均数～数十kbである場合には、そのまま電気泳動による分画を行うことができるため、第二の制限酵素による処理を行う必要はない。

【0035】

(4) 一次元分画

制限酵素Bによる処理を施したのち、各断片について一次元分画を行う（図1(5)）。一次元分画は、直径3～4.5mm、長さ60cm程度の細長いキャピラリー状のチューブを用いることにより行う。上記(4)の項において処理された断片をチューブの原点に注入し、分画する。一次元分画は、0.7～1.0%程度のアガロース電気泳動が好ましく、5%シュクロースの存在下、室温（より好ましくは22～26℃）で20～48時間行う。

【0036】

なお、一次元分画工程により、制限酵素Bによる切断断片は、原点（一次元分画の開始点）から末端部に向かってDNA断片の長さが順に短くなるように泳動される（図1(5)の拡大図）。例えば、図1(5)の拡大図において長さの長い断片1は原点に近く、長さの短い断片2は原点から離れるように泳動する。従って、原点からの位置は、断片の長さを反映することとなる。

【0037】

(5) 第三の制限酵素処理

分画終了後、チューブを第三の制限酵素溶液に浸して一次元分画産物について制限酵素処理を行う。第三の制限酵素は、制限酵素A及びBよりも切断頻度の高い酵素であって、切断したときに生じる平均断片長が数百bp、例えば300bp程度の間隔で存在する制限酵素部位を認識するものである（図1において「C」の位置を切断する酵素であり、制限酵素Cとする。）。

制限酵素Cには4～6塩基認識ものを用いることができ、例えばMboI、HinfI等を採用することができる。

【0038】

(6) 二次元分画

制限酵素Cで断片を処理すると、制限酵素認識部位AとBとで挟まれた断片（A-B断片という）は制限酵素認識部位AとCとで挟まれた断片（A-C断片という）、及び制限酵素認識部位CとBとで挟まれた断片（B-C断片という）に切断され、得られるDNA断片の平均鎖長はそれぞれ数百bp以下になる。

【0039】

これらの断片について二次元分画を行う。二次元分画法としては、例えば、5%

ポリアクリルアミドゲル電気泳動 (2v/cmで20~24時間の条件) による方法等が挙げられる。または、6M尿素を含む5%ポリアクリルアミドゲル電気泳動を行うことによって、より正確なパターンを得ることができる。

【0040】

(7) スポットの検出

スポットの検出は、標識物質の種類に応じた手法を採用することにより行うことができる。例えば、標識物質として³²Pを用いた場合はオートラジオグラフィによる検出が挙げられ、蛍光色素を用いた場合は蛍光イメージアナライザー(例えばFMBIO II Multi-View, TaKaRa)による検出が挙げられる。なお、二次元分画後のB-C断片は標識されていないため、スポットは生じない。

【0041】

こうして得られた各スポットの位置を原点からX、Y方向の距離 (X1 , Y1)、(X2 , Y2)、... (Xn , Yn) で表すと、このX座標は制限酵素Aの認識部位から制限酵素Bの認識部位まで (A-B断片) の距離を、Y座標は制限酵素Aの認識部位から制限酵素Cの認識部位まで (A-C断片) の距離を反映する (図1(6))。従って、スポットのパターンを解析することにより、ゲノムの変異箇所を特定したり、DNA修飾による変異を検出することができる。例えば、スポットの濃さは検出された断片のコピー数を反映するので特定領域の重複であることがわかり、図1(6)のスポットのパターンにおいて二次元分画後の位置 (X2 , Y2) がずれた場合又は消滅した場合は、DNA断片の塩基配列の一部が変異 (欠失、置換又は付加等) したことがわかる。

【0042】

次に、上述したように得られた解析対象のゲノムDNAの二次元電気泳動パターンを解析するプログラムについて説明する。図3に、このプログラムを実行するコンピュータ装置を含むシステムの構成を示すブロック図を示す。また、図4に、このシステムを用いたプログラムの具体的な処理を示すフローチャートを示す。

【0043】

本プログラムを実行するコンピュータは、全体の動作を制御するCPU501と、必

要なプログラムデータ等を一時的に格納するRAM503と、操作者が各種情報を入力する入力部504と、インターネットを介して外部との間でデータの送受信を行う通信モデム等からなる送信／受信部505と、表示するデータ等を出力する出力部505と、データやプログラム等が記録されているHDD507と、CD-ROM509に書き込まれたデータを読み取るCD-ROMドライブ508とを備えている。また、本システムにおいては、インターネット等の通信回線網を介してコンピュータと接続されたデータベース510を備えている。なお、本プログラムは、テープ状又はディスク状磁気記録媒体、光学記録媒体等、一般にコンピュータの外部記憶デバイスに使用される情報記録媒体であればいかなる記録媒体にも記録することができる。

【0044】

CPU501は、RAM503、HDD507又はCD-ROM509に記録されているプログラムに従って、システム全体を制御して後述する処理を実行する。入力部504は、例えば、スキャナーやマウス、キーボード等であり、処理を実行する上で必要な条件等を入力するとき等に操作される。出力部506は、例えば、ディスプレイやプリンター等であり、CPU501の制御によって処理されたデータ等を表示することができる。送信／受信部505は、CPU501の制御によってデータベース510に記録されたデータをRAM503やHDD507等に取り込むことができる。

【0045】

この処理では、図4に示すように、ステップS1と、ステップS2及びステップS3とを平行して処理することができる。この処理においては、ステップS1をステップS2及びステップS3に先立って行っても良いし、ステップS2及びステップS3をステップS1に先立って行っても良い。

【0046】

ステップS1では、上述したように解析対象のゲノムDNAから得られた二次元電気泳動パターンを入力する。具体的には、「(7)スポットの検出」で得られた各スポットの位置 (X_n , Y_n) をデジタルデータとして入力することができる。また、二次元電気泳動パターンをスキャナー等により画像データとして入力することもできる。

【0047】



ステップS2では、データベース510に記録されているゲノム塩基配列情報を、インターネット等の通信回線網を介して取得する。ここで、ゲノム塩基配列情報は、解析対象となるゲノムDNAの野生型として登録されている塩基配列である。データベース510としては、例えば、Genbank、EMBL、DDBJ、TAIR及びKAOS等を挙げることができる。なお、本プログラムは、ゲノム塩基配列情報としてデータベース510に記録されているものに限定されず、例えば、HDD507、CD-ROM509及びその他各種情報記録媒体等に記録されているゲノム塩基配列情報を使用することができる。

【0048】

ステップS3では、データベース510から得たゲノム塩基配列情報を用いてコントロールの二次元電気泳動パターンを作成する。コントロールの二次元電気泳動パターンを作成する際には、先ず、ゲノム塩基配列情報から制限酵素Aの認識配列及び制限酵素Bの認識配列を検索し、A-A断片及びA-B断片を予測する。次に、解析対象の二次元電気泳動パターンがアダプターを使用した場合には、予測したA-A断片及びA-B断片にアダプターの塩基配列数を加算する。これにより、「(3) 第二の制限酵素処理」で得られた複数のA-A断片及びA-B断片に対応する、複数の推定A-A断片及び推定A-B断片を予測することができる。

【0049】

次に、複数の推定A-A断片及び推定A-B断片の塩基配列数及び推定される等電点等に基づいて、「(4) 一次元分画」を行った場合に推定される泳動パターンを予測する。この泳動パターンは、Southern, E.M. (Measurement of DNA Length by Gel Electrophoresis, Analytical Biochemistry:100,319-323,1979) により報告された方法により算出することができる。なお、泳動パターンは、ゲルの濃度及び種類、泳動を行う電圧、泳動時間及び温度等の諸条件をパラメータとして入力することによって推定することができる。言い換えると、本プログラムでは、これらの諸条件をパラメータとして泳動パターンを推定できるような補正値を予め備えることが好ましい。

【0050】

次に、推定A-A断片及び推定A-B断片における制限酵素Cの認識配列を検索し、A



-C断片を予測する。これにより、「(5) 第三の制限酵素処理」で得られた複数のA-C断片に対応する、複数の推定A-C断片を予測することができる。

次に、複数の推定A-C断片の塩基配列数及び推定される等電点等に基づいて、「(6) 二次元分画」を行った場合に推定される二次元電気泳動パターンを予測する。二次元電気泳動パターンは、上述した「泳動パターン」を予測する場合と同様に、ゲルの種類、泳動を行う電圧、泳動時間及び温度等の諸条件をパラメータとして入力することによって推定することができる。言い換えると、本プログラムでは、これらの諸条件をパラメータとして二次元電気泳動パターンを推定できるような補正値を予め備えることが好ましい。

【0051】

本プログラムでは、得られたコントロールの二次元電気泳動パターンに基づいて、予測されるスポットの位置を推定座標 (X_n , Y_n) として得ることができる。このとき、 X 座標 (X_n) は、アダプターの塩基配列数を加算したA-B断片のサイズ分画位置を示し、 Y 座標 (Y_n) は、アダプターの塩基配列数を加算したA-C断片のサイズ分画位置を示す。また、このステップS3では、得られた推定座標 (X_n , Y_n) に基づいて、コントロールの二次元電気泳動パターンを画像データとして得ることもできる。

【0052】

なお、上述した説明では、推定A-B断片に基づいて泳動パターンを作成したが、解析対象の二次元電気泳動パターンを作成するに際して「(3) 第二の制限酵素処理」を行わない場合には、制限酵素A処理によって得られると推定された推定A断片に基づいて泳動パターンを作成すればよい。また、解析対象の二次元電気泳動パターンを作成するに際してアダプターを付加していない場合には、上述した処理においてアダプターの塩基配列数を考慮しなくて良い。

【0053】

さらに、コントロールの二次元電気泳動パターンは、ゲノム塩基配列情報における所定の領域を排除したものとして作成されることが好ましい。排除される領域としては、ゲノムDNA端部のテロメア領域や、ゲノムDNAの動原体付近等の高度反復塩基配列を挙げることができる。具体的には、第三の制限酵素処理前に、一

方の端部のみに制限酵素Aの認識配列を有する断片は、ゲノムDNA端部のテロメア領域由来或いは塩基配列決定の際のギャップ領域由来であるとしてスポットとして同定しない。

【0054】

次に、ステップS4では、ゲノム塩基配列情報に基づいて得られたコントロールの二次元電気泳動パターンと、解析対象の二次元電気泳動パターンとを比較する。具体的には、ステップS1で入力したスポットの座標 (X_n , Y_n) と、ステップS3で得られたスポットの推定座標 (X_n , Y_n) とを比較する。また、ステップS4では、ステップS1で得られた二次元電気泳動パターンの画像データと、ステップS3で得られたコントロールの二次元電気泳動パターンの画像データと比較することもできる。

【0055】

次に、ステップS5では、ステップS4で行った比較の結果を、例えばディスプレイ等に表示する。比較結果としては、コントロールの二次元電気泳動パターンにおいて、解析対象の二次元電気泳動パターンと一致しないスポットを同定し、当該スポットの座標及び塩基配列情報を表示することができる。コントロールの二次元電気泳動パターンは、データベース510から取得したゲノム塩基配列情報に基づいて作成されているため、各スポットの塩基配列情報を対応させることができる。

【0056】

特に、高等生物等の染色体構造を有する細胞を解析対象とする場合、ゲノムDNA端部のテロメア領域や、塩基配列決定の際に不明瞭なデータを示す動原体付近等の高度反復塩基配列を比較対照とすると、ノイズの原因となってしまう。そこで、ゲノムDNA端部のテロメア領域や、ゲノムDNAの動原体付近等の高度反復塩基配列を除外したコントロールの二次元電気泳動パターンを使用することによって、ノイズとなるようなスポットについて比較することを防止できる。一方で、最初の段階において、各クローンの両端をマーキングしておくことで、パターン中出现したノイズを明確に識別できるようにしておく。これらにより、ノイズ成分を低減してより精度の高い比較を行うことができる。

【 0 0 5 7 】

以上説明したように、本発明に係るプログラムによれば、データベース510等に記録されたゲノム塩基配列情報を用いてコントロールの二次元電気泳動パターンを作成し、このコントロールの二次元電気泳動パターンを用いて解析対象の二次元電気泳動パターンを解析する。これにより、このプログラムによれば、ゲノムDNAから得られた二次元電気泳動パターンを解析するに際して、野生型の細胞株から抽出したゲノムDNAを用いてコントロールとなる二次元電気泳動パターンを作成する必要がある。また、本プログラムによれば、多大な労力と煩雑な技術の必要性もなくなる。したがって、このプログラムは、解析対象のゲノムDNAを非常に簡便に且つ迅速に解析することができる。

【 0 0 5 8 】

特に、本プログラムでは、物理的変異原によって塩基の欠失等の変異が誘発されたゲノムDNAを使用することが好ましい。すなわち、本プログラムでは、野生型のゲノムDNAと比較して長さが変化しているようなゲノムDNAを解析対象とすることが好ましい。具体的には、重イオンビームを照射することによって突然変異を誘発するような手法（特開平9-28220号公報参照）を用いて、突然変異が導入されたゲノムDNAを解析対象とする。

【 0 0 5 9 】

また、特に、本プログラムを適用して植物細胞におけるゲノムDNAを解析する場合、植物ゲノムDNAにおけるシトシンがメチル化されていることが多いため、第一の制限酵素としてメチル化非感受性の制限酵素を用いることが好ましい。すなわち、植物ゲノムDNAを解析対象とする場合には、メチル化非感受性制限酵素を使用することによって、メチル化の影響を受けることなく確実に切断することができる。第一の制限酵素処理に使用するメチル化非感受性制限酵素としては、例えば、AccIII、PacIを挙げることができる。また、第二の制限酵素処理に使用するメチル化非感受性制限酵素としては、4塩基認識のMboIを挙げることができる。これらの制限酵素を使用することによって、植物ゲノムDNAを解析対象とする場合であっても、メチル化の影響を受けることがなく、多くのスポットを得ることができ、より精密な解析を行うことができる。

なお、ゲノムDNAを解析する場合、メチル化感受性制限酵素を用いてもよい。
 この場合、第一の制限酵素としてメチル化感受性制限酵素を使用すると、DNAのメチル化による後生的変異領域の検出及び解析をすることができる。

【0060】

【実施例】

以下、実施例を用いて本発明をより詳細に説明するが、本発明の技術的範囲はこの実施例に限定されるものではない。

【0061】

実施例1：シロイヌナズナ葉緑体DNA

＜コントロール二次元電気泳動パターンの作成＞

本例では、シロイヌナズナの葉緑体DNAを解析対象とした例である。

【0062】

シロイヌナズナ (Arabidopsis thaliana, (Columbia系統)) の葉緑体DNAは、154,478bpの環状ゲノムDNAを有するものとして塩基配列が既に決定されており (Sato, S. et al. 1999, DNA RESEARCH 6,238-290)、アクセッションナンバーAP000423としてDDBJからその塩基配列データを得ることができる。この塩基配列データに基づいて前述のプログラム工程を経てコントロールの二次元電気泳動パターンを作成した。

【0063】

本例では、第一の制限酵素としてAccIII、第二の制限酵素を使用せず、第三の制限酵素としてMboIを用いたときの二次元電気泳動パターンを作成した。これらAccIII及びMboIはともに、シトシンがメチル化してなる5メチルシトシン (5mC) に対しては非感受性である。得られたコントロールの二次元電気泳動パターンを図5に示す。図5中、塩基数の表示は実際に分子量マーカを泳動したときの移動度を示し、横軸が一次元分画、縦軸が二次元分画の展開を示している。

【0064】

本例で作成したコントロールの二次元電気泳動パターン (図5) におけるスポットCP3-aの塩基配列を配列番号2に示し、スポットCP6-bの塩基配列を配列番号3に示し、スポットCP16-aの塩基配列を配列番号4に示し、スポットCP16-bの塩

基配列を配列番号5に示し、スポットCP19-aの塩基配列を配列番号6に示し、スポットCP19-bの塩基配列を配列番号7に示し、スポットCP22-aの塩基配列を配列番号8に示し、スポットCP25-bの塩基配列を配列番号9に示し、スポットCP28-aの塩基配列を配列番号10に示し、スポットCP28-bの塩基配列を配列番号11に示し、スポットCP310-aの塩基配列を配列番号12に示し、スポットCP310-bの塩基配列を配列番号13に示し、スポットCP15-aの塩基配列を配列番号14に示し、スポットCP15-bの塩基配列を配列番号15に示す。

【0065】

＜検出対象の二次元電気泳動パターンの作成＞

シロイヌナズナ (Columbia系統) の葉緑体DNAを解析対象として、AccIIIとMboIを用いて、実際に二次元電気泳動パターンを以下のように作成した。まず、抽出した葉緑体DNAをAccIIIで消化し、シークナーゼを用いて $[\alpha\text{-}^3\text{ }^2\text{P}]\text{dCTP}$ や $[\alpha\text{-}^3\text{ }^2\text{P}]\text{dGTP}$ で30分間標識反応した。その後、0.8%アガロースゲル電気泳動を2 V/cmで48時間、室温(24℃)という条件で行った。さらに、ゲル中で第三の制限酵素としてMboIを1チューブあたり500unitsで37℃、2時間行なった。その後、5%アクリルアミド電気泳動により2次元分画を行なった。5%アクリルアミド電気泳動の条件は、2 V/cm、22時間、室温(24℃)とした。

二次元電気泳動パターンは、ゲルを乾燥後、14日間、-80℃で露光したのち、現像した。葉緑体DNAを解析対象とした二次元電気泳動パターンを図6に示す。なお、図6中、白抜きの三角印は、観察されるスポットの位置を示している。

【0066】

＜両パターンの比較からのスポット同定＞

図5に示したコントロールの二次元電気泳動パターンと、図6に示した解析対象の二次元電気泳動パターンとを比較した。これら図5及び図6から明らかなように、各二次元電気泳動パターンにおけるスポットの位置がほぼ一致していることが判る。

【0067】

図6に示したスポットのうち、図5中CP3-a(配列番号2)及びCP15-b(配列番号15)で示す推定スポットと対応するスポットからDNA断片を抽出し、クロ

ーン化して塩基配列を決定した。その結果、データベースから得られるゲノム配列情報におけるCP3-a及びCP15-bの塩基配列の情報とそれぞれ完全に一致していた。

このことより、ゲノム塩基配列情報を用いて作成したコントロールの二次元電気泳動パターンを用いて、実際の解析対象の二次元電気泳動パターンを解析できることを明確に示すことができた。

【0068】

また、図6に示した解析対象の二次元電気泳動パターンから、葉緑体DNAを用いると、スポットが強いシグナル強度を示すことが判る。スポットが強いシグナル強度を示すのは、葉緑体DNAのコピー数が多数あるためである。このことから、解析対象の二次元電気泳動パターンにおいては、解析対象のゲノムDNAの量に比例して強いシグナル強度で表示できることが判った。

【0069】

実施例2：シロイヌナズナ核ゲノムDNA

<コントロール二次元電気泳動パターンの作成>

本例では、シロイヌナズナの核ゲノムDNAを解析対象とした例である。シロイヌナズナ (*Arabidopsis thaliana*, (Columbia系統)) の核ゲノムDNAは、約130 Mbとして、一部ギャップ領域（未決定領域）があるもののほぼ全体の塩基配列が既に決定されている (The Arabidopsis Genome Initiative, Nature, 408 796-815 (2000)参照、The Arabidopsis Information resource(TAIRと称される、www.arabidopsis.org/home.html)において開示されている)。この塩基配列データに基づいて前述のプログラム工程を経てコントロールの二次元電気泳動パターンを作成した。

【0070】

本例では、第一の制限酵素としてNotI、第二の制限酵素としてEcoRV、第三の制限酵素としてMboIを用いたときの二次元電気泳動パターンを作成した。これらEcoRV及びMboIはともにシトシンがメチル化してなる5メチルシトシン (5mC) に対しては非感受性であり、NotIは感受性である。得られたコントロールの二次元電気泳動パターンを図7に示す。図7中、塩基数の表示は実際に分子量マーカー

を泳動したときの移動度を示し、横軸が一次元分画、縦軸が二次元分画の展開を示している。また、図7中、各スポットに対応して記載されている文字列は、ゲノムDNA中のNotIサイトを含む遺伝子座名を示している。左端の1~5の数字は染色体番号を示し、右端の「-a」及び「-b」はそれぞれNotIサイトの上流側及び下流側を示し、それ以外の文字列は塩基配列を決定する際に作成されたクローン名を意味している。例えば、「4FCA1-b」がリンクしたスポットは、第4番染色体に由来し、FCA1クローン中のNotIサイトの下流側のDNA断片を含み、塩基配列が決定されたDNA断片を意味することとなる。なお、図7中、黒塗りの円(●)は第1染色体由来のスポットであり、黒塗り菱形(◆)は第2染色体由来のスポットであり、黒塗りの三角(▲)は第3染色体由来のスポットであり、クロス(×)は第4染色体由来のスポットであり、黒塗り四角(■)は第5染色体由来のスポットである。

【0071】

＜検出対象の二次元電気泳動パターンの作成＞

X線照射によって突然変異を誘発したシロイヌナズナ(Columbia系統)のゲノムDNAを解析対象とし、第一の制限酵素としてNotI、第二の制限酵素としてEcoRV、第三の制限酵素としてMboIを用いた以外は実施例1と同様にして、実際に二次元電気泳動パターンを作成した。得られた二次元電気泳動パターンを図8中「Mutant」に示す。

【0072】

なお、図8において、「Control」はコントロールの二次元電気泳動パターンの一部を抜き出して拡大したものである。これら「Mutant」と「Control」との比較により、X線照射によって4FCA1-bが消失していることが確認された。また、4FCA1-bの両側のNotIサイトのスポットの有無を調べたところ、上流の4T15F16-b、下流の4T16H5-aや4T805-aのスポットを確認することができることから、4FCA1-b及び4T16H5-a間の最長で約2.4Mb程度の欠失が起こっていることが示唆された。この欠失を検証するために、4FCAと4T16H5との間にある4F13C5領域の一部をプローブとするサザン解析を行った。その結果、図8中「Southern analysis」に示すように、この領域の欠失を確認することができた。なお、図8中「Southern anal

ysis」において、CレーンはX線照射を行っていないサンプルであり、MレーンはX線照射を行ったサンプルである。

【0073】

このように、図7に示したコントロールの二次元電気泳動パターンを用いることによって、突然変異体における変異領域を検出できることが明らかになった。すなわち、解析対象の二次元電気泳動パターンとコントロールの二次元電気泳動パターンとを比較することによって、非常に短時間に効率よく突然変異体における変異領域を検出できることが明らかになった。

【0074】

ところで、得られた二次元電気泳動パターン（図8中「Mutant」）においては、図7に示したコントロールの二次元電気泳動パターンにおけるいくつかのスポットを確認することができなかった。これはNotIがメチル化に対して感受性であり、染色体DNAにおけるメチル化の影響によってスポットが出現しないためである。脱メチル化処理によって染色体DNAを低メチル化した場合には、これらスポットを確認することができる。

したがって、図7に示したコントロールの二次元電気泳動パターンを用いることによって、突然変異体における変異領域を検出できるのみならず、DNAメチル化等のDNA修飾も検出できることが明らかになった。

【0075】

【発明の効果】

以上、詳細に説明したように、本発明によれば、解析対象のゲノムDNAを比較するためのコントロールの二次元電気泳動パターンを容易に、且つ、様々な条件で作成することができる。したがって、本発明によれば、ゲノムDNAを用いて作成した二次元電気泳動パターンを容易に解析することができるプログラムを記録した記録媒体を提供することができる。

また、本発明によれば、ゲノム塩基配列情報に基づいてコントロールの二次元電気泳動パターンを作成することによって、迅速に且つ確実な解析を行うことができるゲノムDNA解析方法を提供することができる。

【0076】

【配列表】

SEQUENCE LISTING

<110> RIKEN

<120> A recording media having analysis program for genomic
DNA

<130> RJH12-138

<140>

<141>

<150> JP2000-327632

<151> 2000-10-26

<160> 15

<170> PatentIn Ver. 2.0

<210> 1

<211> 12

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic DNA

<220>

<221> modified base

<222> 4, 5, 6, 7, 8 and 9

<223> n represents a, g, c or t

<400> 1

ccannnnnnt gg

12

<210> 2

<211> 207

<212> DNA

<213> Arabidopsis thaliana (Columbia)

<400> 2

tccggataat atccataaat gacttgTTTT tggtaagagg tggacattac tatatgtaaa 60
 ttcgggtgca tgggacacat cagtactcca gtgcatttcg ccctctgagt cagaataaat 120
 atattttcta accctctctt taaaatgaaa agtgtatgtt ccctcgcgaa tctcagcaat 180
 cacttgTtct gattccacat attgate 207

<210> 3

<211> 163

<212> DNA

<213> Arabidopsis thaliana (Columbia)

<400> 3

gatcttttat actgtgtccg attccaaagt tcgttctata catatgaccc gcaatgagga 60
 aaagaattgc gatagctaaa tgatgatgtg ccatatcggt tagceataaa ctttgcgttt 120
 gtggatggaa tcccccaaga agggtagaa tggcagttcc gga 163

<210> 4

<211> 131

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 4

tccggaatat gagtgtgtga cttgttagaa ttgaccctat ggatagtaca gagaatgggg 60
tctgtcatct ttatcaagat ggttttactt cgtcggatat tcattcgagt atctggagca 120
cgaaatagat c 131

<210> 5

<211> 134

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 5

gaccttcct ttattcaaac ggaaggaaga gagatagaat cagaccgatt ccctaaatac 60
ctttctggct attcctcaat gccccggcta ttcacggaac gtgagaagcg aatgaataat 120
catctgcttc cgga 134

<210> 6

<211> 189

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 6

tccggaggat gccttatata tatattaata tatattatat caaaaagatg gacaatcaaa 60
tctatttctc gattcaatag aagtccaacc aaagaggtga atagggtccc aaataacgag 120
agatatgtaa aaagtaggtc agatttcgcc tattcctaata cctaaatgga atgtaacgac 180

gtagggatc

189

<210> 7

<211> 232

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 7

gatcagccac actgggactg agacacggcc cagactccta cgggaggcag cagtggggaa 60
 ttttccgcaa tgggcgaaag cctgacggag caatgccgcg tggaggtaga aggcctacgg 120
 gtcctgaact tcttttccca gagaagaagc aatgacggta tctggggaat aagcatcggc 180
 taactctgtg ccagcagccg cggtaatata gaggatgcaa gcgttatccg ga 232

<210> 8

<211> 597

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 8

tccggagatt cccgaatagg ttaacctttt gaactgctgc tgaatccatg ggcaggcaag 60
 agacaacctg gcgaactgaa acatcttagt agccagagga aaagaaagca aaagcgattc 120
 ccgtagtagc ggcgagcgaa atgggagcag cctaaaccgt gaaaacgggg ttgtgggaga 180
 gcaaaaaaag cgtcgtgctg ctaggcgaag cgggtggagt cgcacccta gatggcgaga 240
 gtccagtagc cgaaagcatc actagcttat gctctgacct gagtagcatg gggcacgtgg 300
 aatcccgtgt gaatcagcaa ggaccacctt gcaaggctaa atactcctgg gtgaccgata 360
 gcgaagtagt accgtgaggg aagggtgaaa agaaccacca tcggggagtg aaatagaaca 420
 tgaaaccgta agctcccaag cagtgggagg agccctgggc tetgacegeg tgeetgttga 480
 agaatgagcc ggcgactcat aggcagtggc ttggttaagg gaaccaccg gagccgtagc 540
 gaaagcgagt cttcataggg caattgtcac tgcttatgga cccgaacctg ggtgatc 597

<210> 9

<211> 597

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 9

gatcaccag gttcgggtcc ataagcagt acaattgccc tatgaagact cgctttcgct 60
acggctccgg tgggttcct taaccaagcc actgcctatg agtcgccggc tcattcttca 120
acaggcacgc ggtcagagcc cagggtcct cccactgctt gggagcttac ggtttcatgt 180
tctatttcac tccccgatgg gggttcttt caccctccc tcacgtact acttcgctat 240
cggtcacca ggagtattta gccttgcaag gtggtccttg ctgattcaca cgggattcca 300
cgtgccccat gctactcggg tcagagcata agctagtgat gctttcggct actggactct 360
cgccatctag ggtgcggcac tccaccgctt cgcctagcag cagcagctt ttttgctct 420
cccacaaccc cgttttcacg gtttaggctg ctcccatttc gctcgccgct actacgggaa 480
tcgcttttgc tttctttcc tctggctact aagatgtttc agttcgccag gttgtctctt 540
gcctgccccat ggattcagca gcagttcaaa aggttaacct attcggaat ctccgga 597

<210> 10

<211> 232

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 10

tccggataac gcttgcatec tctgtattac cgcggctgct ggcacagagt tagccgatgc 60
ttattcecca gataccgtca ttgcttctt tctgggaaaa gaagttcagg acccgtaggc 120
cttctacctc cagcggcat tgcctcgta ggctttcgcc cattgeggaa aatteccac 180
tgctgcctcc cgtaggagtc tgggccgtgt ctacgtccca gtgtggctga tc 232

<210> 11

<211> 190

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 11

gatccctacg tcgttacatt ccatttagga ttaggaatag gcgaaatctg acctactttt 60
 tacatatctc tcgttatttg ggaccctatt cacctctttg gttggacttc tattgaatcg 120
 agaaatagat ttgattgtcc atctttttga tataatatat attaatatat atataaggca 180
 tccttccgga 190

<210> 12

<211> 134

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 12

tccggaagca gatgattatt cattcgcttc tcacgttccg tgaatagccg gggcattgag 60
 gaatagccag aaaggtattt agggaatcgg tctgattcta tctctcttcc ttccgtttga 120
 ataaaggaag gatc 134

<210> 13

<211> 215

<212> DNA

<213> *Arabidopsis thaliana* (Columbia)

<400> 13

gatcttgtgg gaagagtgtt gccaaattta tactgtccgc tatctcatag gtggctacgg 60
 gccgcaccaa aacaaaaaac ttttcttgg ttggtgtaat ccgttggaca taaatccaat 120

ttttaaat tttggattcc ttcgagtttt tttttcctct tcctggcggc atcaagatgc 180
cactgtgtcg ggatatttta tctgtcttgt ccgga 215

<210> 14

<211> 618

<212> DNA

<213> Arabidopsis thaliana (Columbia)

<400> 14

tccggaggag cacgaatgca agaaggaagt ttaagtttga tgcaaatggc taaaatttct 60
tcggttttat gtgattatca atcaagtaaa aagttattct atatatcaat tcttacatct 120
cctactaccg gtggagtgc agctagtttt ggtatgttgg gggatatcat tattgccgaa 180
ccctatgcct atattgcatt tgcgggtaaa agagtaattg aacaaacatt gaaaaagcc 240
gtgcctgaag gttcacaagc ggctgaatct ttattacgta agggcttatt ggatgcaatt 300
gtaccacgta atcttttaaa aggtgttctg agcgagttat ttcagctcca tgcttttttt 360
cctttgaaca caaattaaat aaaatagaac ggtagtttta tcagaattaa acgaaaaccc 420
agaaaaatgc atttttcttt caaatcattt tttttatcg atattcttgt ttactactca 480
gtaaacctct atcaacaagc taaaaagtga attttttgg gggggaagtt caaattagac 540
tagacaaaca aaaaaagtt cattttcctc ccttgcttgc atatgtatag ataattcaaa 600
tatagataga tgcagatc 618

<210> 15

<211> 323

<212> DNA

<213> Arabidopsis thaliana (Columbia)

<400> 15

gatctgggcc aagaagcact actgtaggga agttattgaa accattgaat tccgaatatg 60
gtaaagtagc tcccggatgg ggaacgaccc ctttgatggg tgttgcaatg gctctatttg 120

cggtattcct atctattatt ttggagattt ataattcttc tgttctactg gatggaattt 180
 cagtgaatta gactgagaag aatcttgaag ttctagcttt tagctcgata caaaaaagta 240
 aagtatgcag gtctaacaat tttagcctat tctcctttgg tagttcgacc gcgaaatttt 300
 tttctgcatt gtatatttcc gga 323

【 0 0 7 7 】

【配列表フリーテキスト】

配列番号 1 : nは、a、g、c又はtを表す（存在位置：第4塩基～第9塩基）

【図面の簡単な説明】

【図 1】

解析対象のゲノムDNAを用いて二次元電気泳動パターンを作成する工程を示す図である。

【図 2】

アダプターを用いて標識を行う工程を示す図である。

【図 3】

本発明に係るプログラムを記録した記録媒体を実行するシステムを概略的に示す構成図である。

【図 4】

プログラムを実行する処理を示すフローチャートである。

【図 5】

シロイヌナズナの葉緑体DNAの塩基配列情報に基づいて作成したコントロールの二次元電気泳動パターンを示す図である。

【図 6】

シロイヌナズナの葉緑体DNAを用いて作成した解析対象の二次元電気泳動パターンを示す写真である。

【図 7】

シロイヌナズナの染色体DNAの塩基配列情報に基づいて作成したコントロールの二次元電気泳動パターンを示す図である。

【図 8】

コントロールの二次元電気泳動パターンを用いて、実際の解析を行った例を示

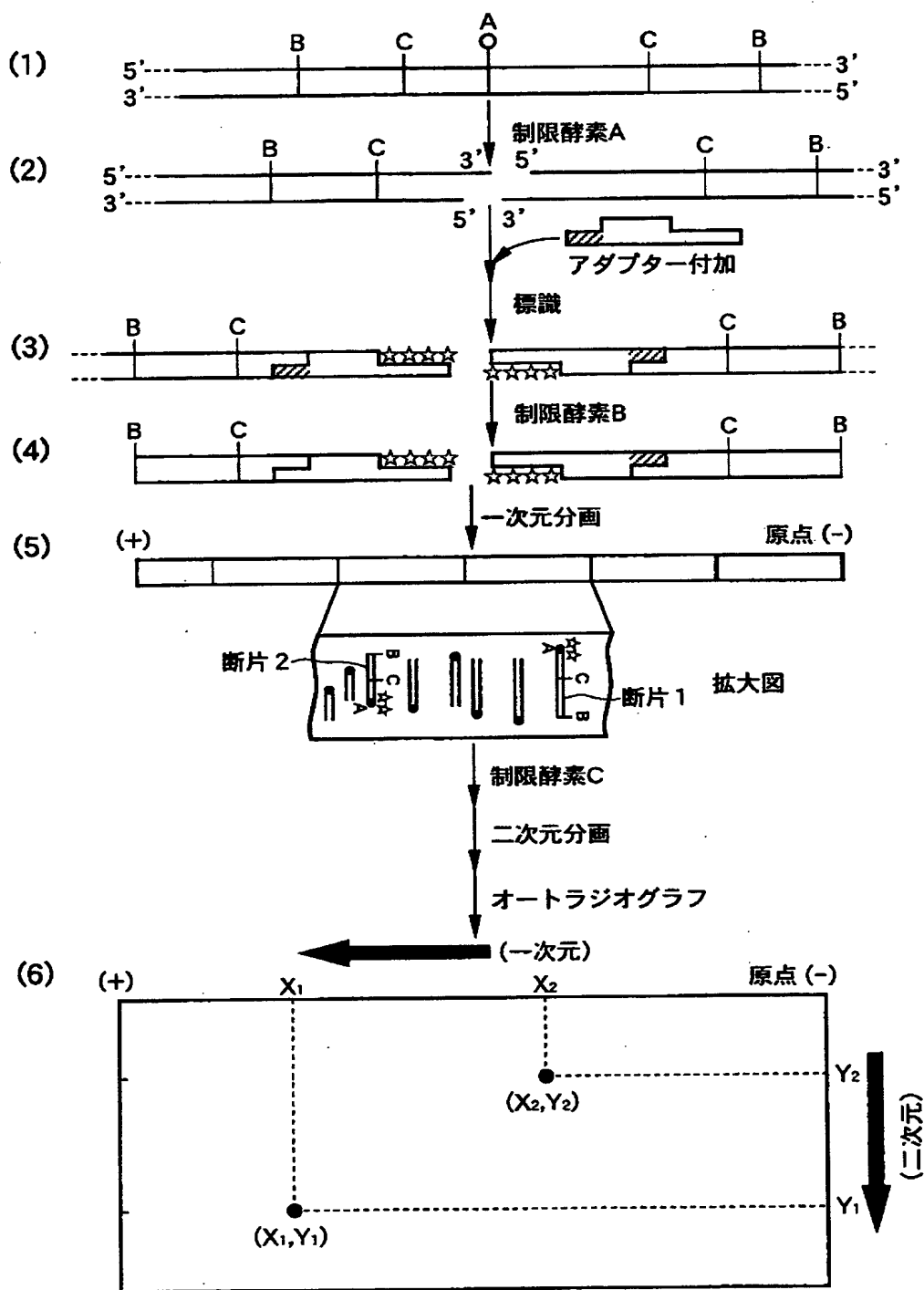
す模式図である。

【符号の説明】

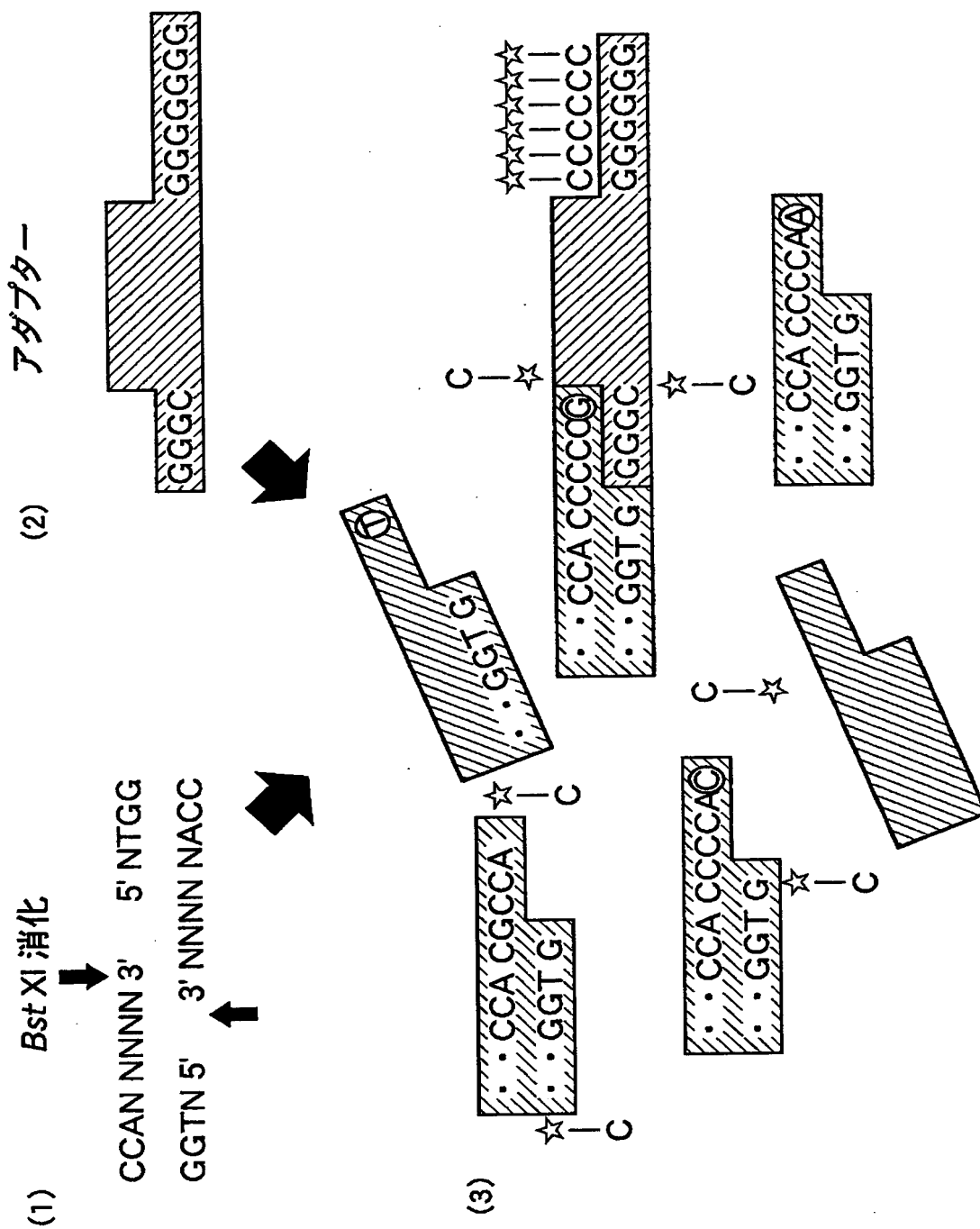
501 CPU、503 RAM、504 入力部、505 送信/受信部、506 出力部、507 HDD
、508 CD-ROMドライブ、509 CD-ROM、510 データベース

【書類名】 図面

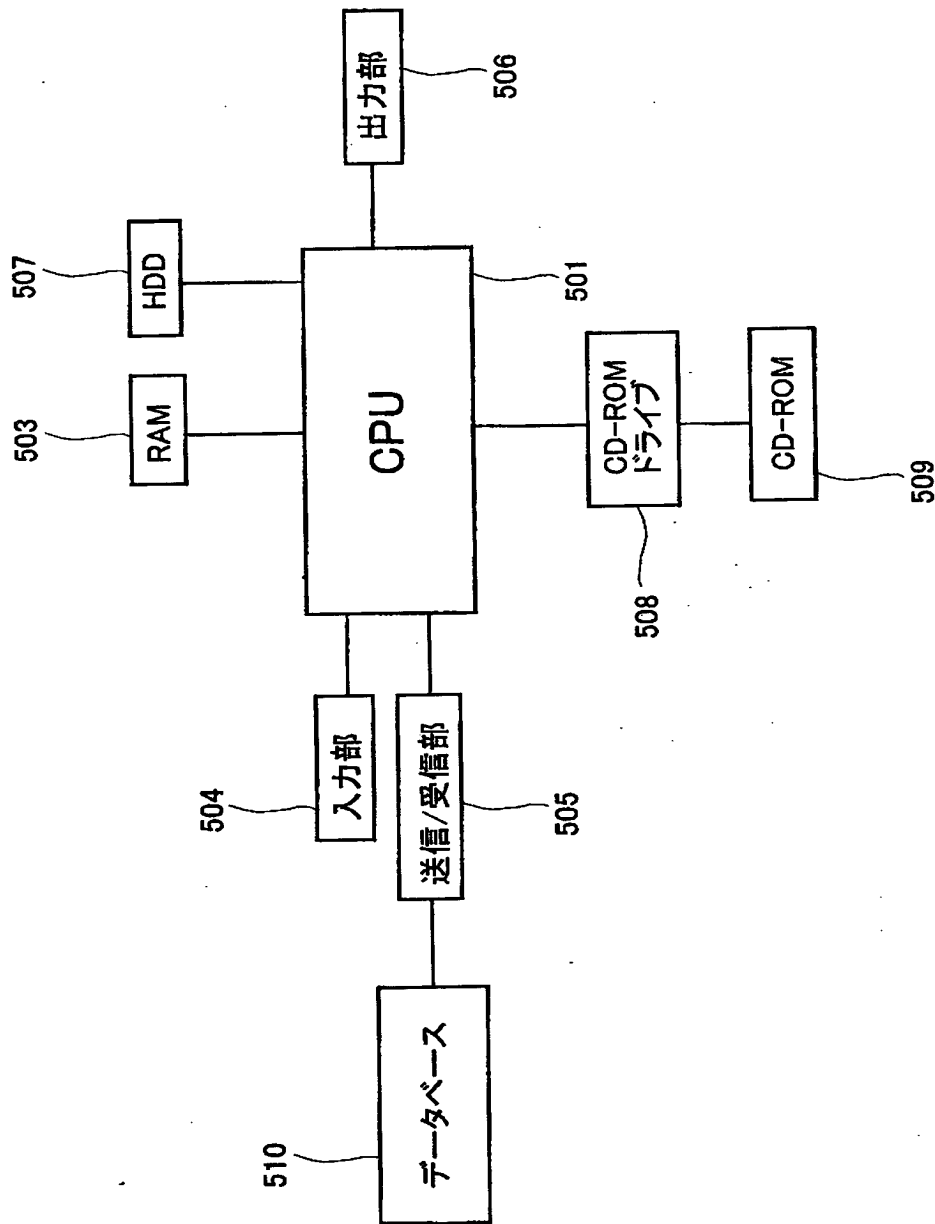
【図 1】



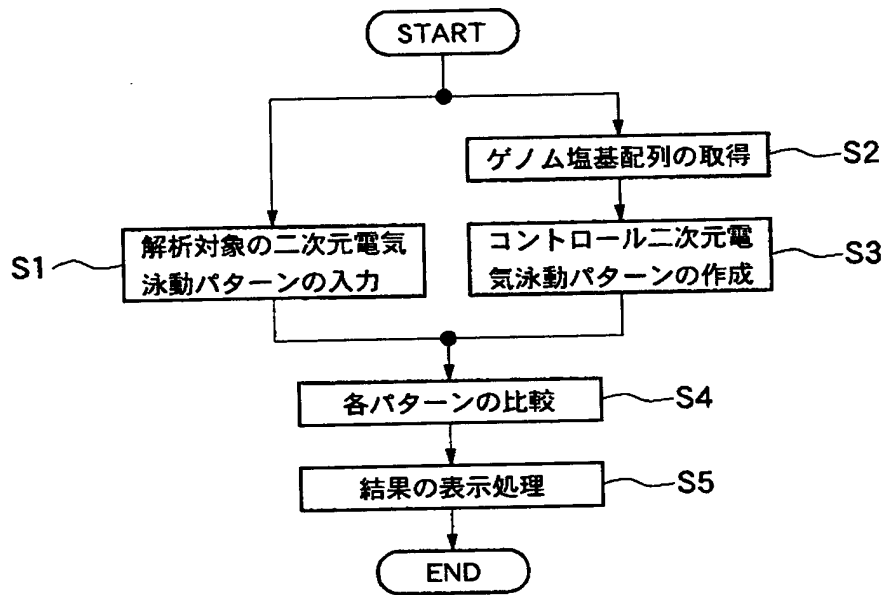
【図 2】



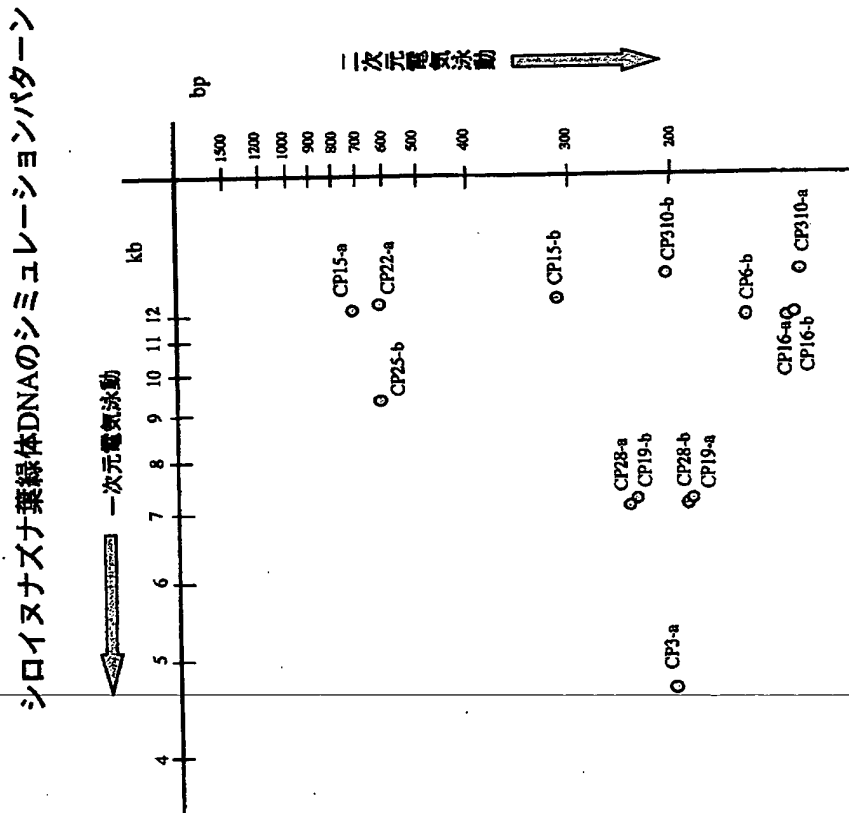
【図 3】



【図 4】

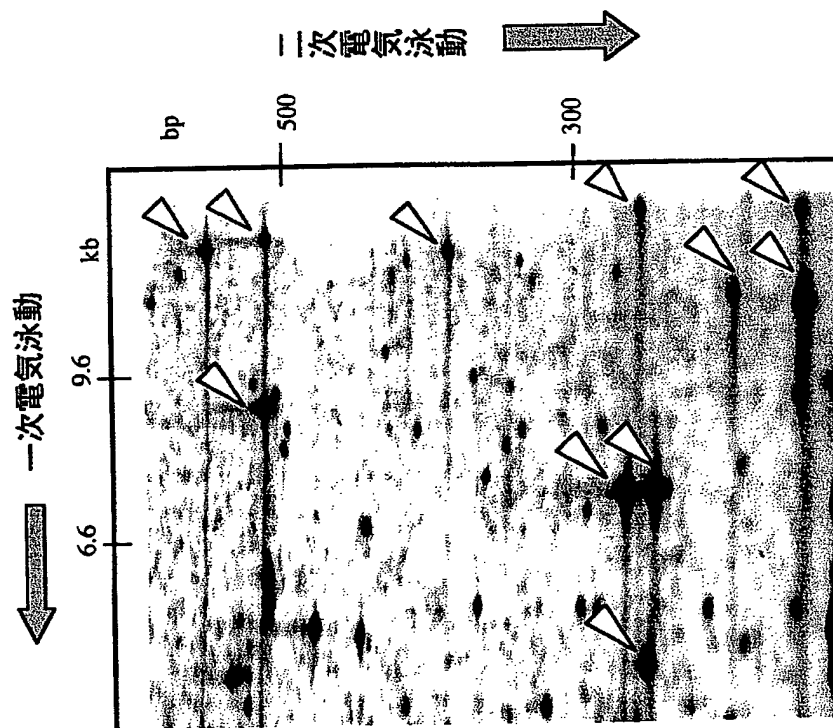


【図 5】

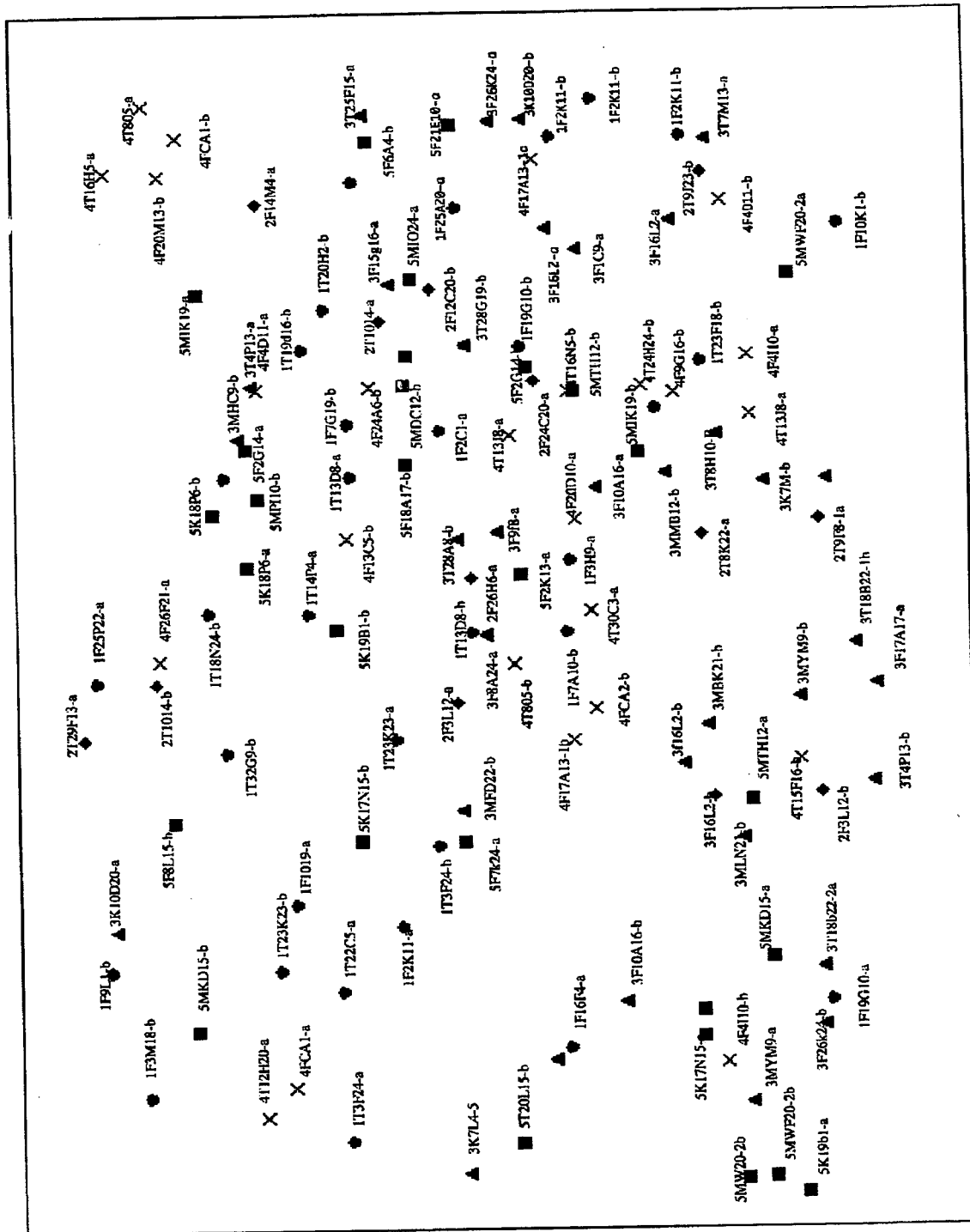


【図 6】

シロイヌナズナ葉緑体DNAのRLGSパターン (AccIII-MboI系)

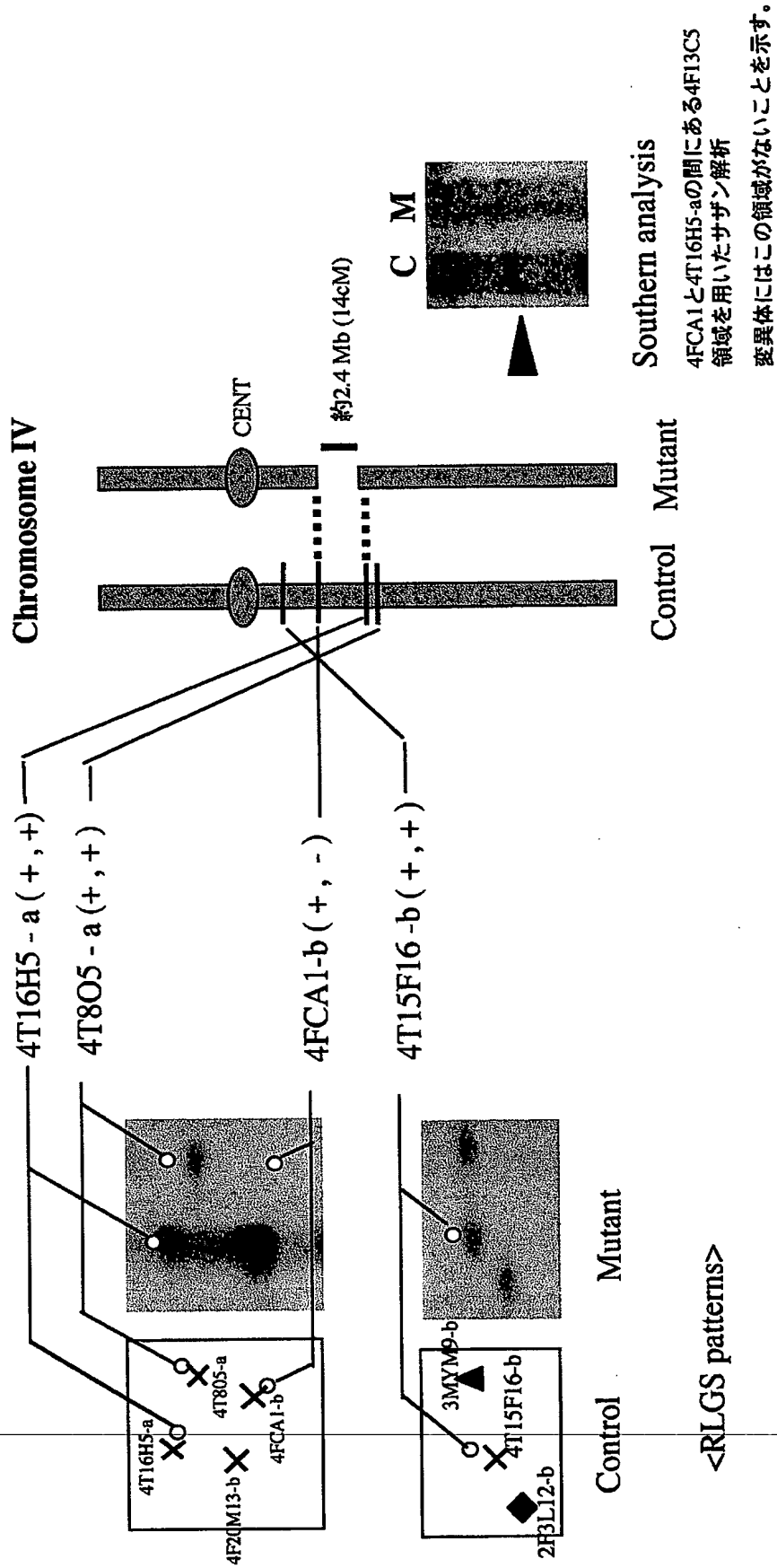


【図7】



【図8】

X線により誘発された突然変異体の解析



【書類名】 要約書

【要約】

【課題】 比較の対象となるコントロールの二次元電気泳動パターンを容易に、且つ、様々な条件で作成し、迅速且つ正確な解析を行う。

【解決手段】 コンピュータに、ゲノム塩基配列情報に基づいて、コントロールの二次元電気泳動パターンを作成する手順と、上記コントロールの二次元電気泳動パターンと、解析対象のゲノムDNAを用いて二次元電気泳動を行って得られた解析対象の二次元電気泳動パターンとを比較する手順とを含み、上記コントロールの二次元電気泳動パターンと上記解析対象の二次元電気泳動パターンとにおけるスポット位置の差異を検出させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体である。

【選択図】 図4

出 願 人 履 歴 情 報

識別番号 [000006792]

1. 変更年月日 1990年 8月28日
[変更理由] 新規登録
住 所 埼玉県和光市広沢2番1号
氏 名 理化学研究所